

# | 000 | manuscripts in practice

Perl, Armenian history, and why Unicode  
rocks

**Where do our histories  
come from?**



SCRIPTORIUM MONK AT WORK. (From *Lacroix*.)



# People aren't machines



- the scriptorium is cold

- the scriptorium is cold
- the food is bad

- the scriptorium is cold
- the food is bad
- the tea ladies are unfriendly



- the scriptorium is cold
- the food is bad
- the tea ladies are unfriendly
- they just don't want to be there anymore

# 600 years later

# 600 years later

- a bunch of error-ridden copies

# 600 years later

- a bunch of error-ridden copies
- or only a few error-ridden copies

# 600 years later

- a bunch of error-ridden copies
- or only a few error-ridden copies
- or only a name check in a book about something else

# Critical editions



Thou craven  
boil-brained  
puttock!



# Apparatus example

## St. Stephen's Church in Nijmegen

Nobilis itaque comes Otto imperio et dominio Novimagensi sibi, ut praefer-  
tur, impignoratis et commissis proinde praeesse cupiens, anno liiii superius 1254  
descripto, mense Iunio, una cum iudice, scabinis ceterisque civibus civitatis  
Novimagensis, pro ipsius et inhabitantium in ea necessitate, commodo et utili-  
5 tate, ut ecclesia eius parochialis extra civitatem sita destrueretur et infra muros  
transferretur ac de novo construeretur, a reverendo patre domino Conrado de  
Hofsteden, archiepiscopo Coloniensi, licentiam, et a venerabilibus dominis de-  
cano et capitulo sanctorum Apostolorum Coloniensi, ipsius ecclesiae ab antiquo  
veris et pacificis patronis, consensum, citra tamen praeiudicium, damnum aut  
10 gravamen iurium et bonorum eorundem, impetravit.

Et exinde liberum locum eiusdem civitatis qui dicitur Hundisbrug, de prae-  
libati Wilhelmi Romanorum regis, ipsius fundi domini, consensu, ad aedifican-  
dum et consecrandum ecclesiam et coemeterium, eisdem decano et capitulo de  
expresso eiusdem civitatis assensu libera contradiderunt voluntate, obligantes  
15 se ipsi comes et civitas dictis decano et capitulo, quod in recompensationem  
illius areae infra castrum et portam, quae fuit dos ecclesiae, in qua plebanus  
habitare solebat—quae tunc per novum fossatum civitatis est destructa—aliam  
aream competentem et ecclesiae novae, ut praefertur, aedificandae satis conti-  
guam, ipsi plebano darent et assignarent. Et desuper apud dictam ecclesiam  
20 sanctorum Apostolorum est littera sigillis ipsorum Ottonis comitis et civitatis  
Novimagensis sigillata.

3 p. 227 R 4 p. 97 N 6 p. 129 D 12 f. 72v M 13 p. 228 R 20 p. 130 D

2 proinde ] primum D 5 ecclesia eius ] ecclesia D: eius eius H extra civitatem om. H

infra ] intra D 6 transferretur ] transferreretur NH 7 Hofsteden ] Hoffstede D: Hoffsteden

H Coloniensi ] Colononiensi H dominis ] viris H 8 Coloniensi ] Coloniae H 10 iurium ]

virium D 11 liberum ] librum H qui ] quae D Hundisbrug ] Hundisburch D: Hunsdisbrug

R 12 regis ] imperatoris D 13 et consecrandum om. H eisdem ] eiusdem D 15 comes ]

comites D dictis om. H 17 tunc ] nunc H 18 ut. . . aedificandae om. H 18–19 contiguam ]

contiguum M 19 apud om. H 20 est ] et H littera ] litteram H 21 Novimagensis ]

Novimagii D sigillata ] sigillis communita H

# Apparatus example

## St. Stephen's Church in Nijmegen

Nobilis itaque comes Otto imperio et dominio Novimagensi sibi, ut praefer-  
tur, impignoratis et commissis proinde praeesse cupiens, anno liiii superius 1254  
descripto, mense lunio, una cum iudice, scabinis ceterisque civibus civitatis  
Novimagensis, pro ipsius et inhabitantium in ea necessitate, commodo et utili-  
5 tate, ut ecclesia eius parochialis extra civitatem sita destrueretur et infra muros  
transferretur ac de novo construeretur, a reverendo patre domino Conrado de  
Hofsteden, archiepiscopo Coloniensi, licentiam, et a venerabilibus dominis de-  
cano et capitulo sanctorum Apostolorum Coloniensi, ipsius ecclesiae ab antiquo  
veris et pacificis patronis, consensum, citra tamen praeiudicium, damnum aut  
10 gravamen iurium et bonorum eorundem, impetravit.

Et exinde liberum locum eiusdem civitatis qui dicitur Hundisbrug, de prae-  
libati Wilhelmi Romanorum regis, ipsius fundi domini, consensu, ad aedifican-  
dum et consecrandum ecclesiam et coemeterium, eisdem decano et capitulo de  
expresso eiusdem civitatis assensu libera contradiderunt voluntate, obligantes  
15 se ipsi comes et civitas dictis decano et capitulo, quod in recompensationem  
illius areae infra castrum et portam, quae fuit dos ecclesiae, in qua plebanus  
habitare solebat—quae tunc per novum fossatum civitatis est destructa—aliam  
aream competentem et ecclesiae novae, ut praefertur, aedificandae satis conti-  
guam, ipsi plebano darent et assignarent. Et desuper apud dictam ecclesiam  
20 sanctorum Apostolorum est littera sigillis ipsorum Ottonis comitis et civitatis  
Novimagensis sigillata.

3 p. 227 R 4 p. 97 N 6 p. 129 D 12 f. 72v M 13 p. 228 R 20 p. 130 D

2 proinde ] primum D 5 ecclesia eius ] ecclesia D: eius eius H extra civitatem om. H  
infra ] intra D 6 transferretur ] transferreretur NH 7 Hofsteden ] Hoffstede D: Hoffsteden  
H Coloniensi ] Colononiensi H dominis ] viris H 8 Coloniensi ] Coloniae H 10 iurium ]  
virium D 11 liberum ] librum H qui ] quae D Hundisbrug ] Hundisburch D: Hunsdisbrug  
R 12 regis ] imperatoris D 13 et consecrandum om. H eisdem ] eiusdem D 15 comes ]  
comites D dictis om. H 17 tunc ] nunc H 18 ut. . . aedificandae om. H 18–19 contiguam ]  
contiguum M 19 apud om. H 20 est ] et H littera ] litteram H 21 Novimagensis ]  
Novimagii D sigillata ] sigillis communita H

# Apparatus example

2 proinde] primum D 5 ecclesia eius] ecclesia D: eius eius H extra civitatem *om.* H  
infra] intra D 6 transferretur] transferreretur NH 7 Hofsteden] Hoffstede D: Hoffsteden  
H Coloniensi] Colononiensi H dominis] viris H 8 Coloniensi] Coloniae H 10 iurium]  
virium D 11 liberum] librum H qui] quae D Hundisbrug] Hundisburch D: Hunsdisbrug  
R 12 regis] imperatoris D 13 et consecrandum *om.* H eisdem] eiusdem D 15 comes]  
comites D dictis *om.* H 17 tunc] nunc H 18 ut...aedificandae *om.* H 18–19 contiguam]  
contiguum M 19 apud *om.* H 20 est] et H littera] litteram H 21 Novimagensis]  
Novimagii D sigillata] sigillis communita H

- **Endless scope for scholarly analysis**

- Endless scope for scholarly analysis
- Gives academics something to do

- Endless scope for scholarly analysis
- Gives academics something to do
- Provides excuses for arguments over port after dinner

# The Chronicle of Matthew



105. *The Third Byzantine Expansion into Armenia, 966–1064*



105. *The Third Byzantine Expansion into Armenia, 966–1064*





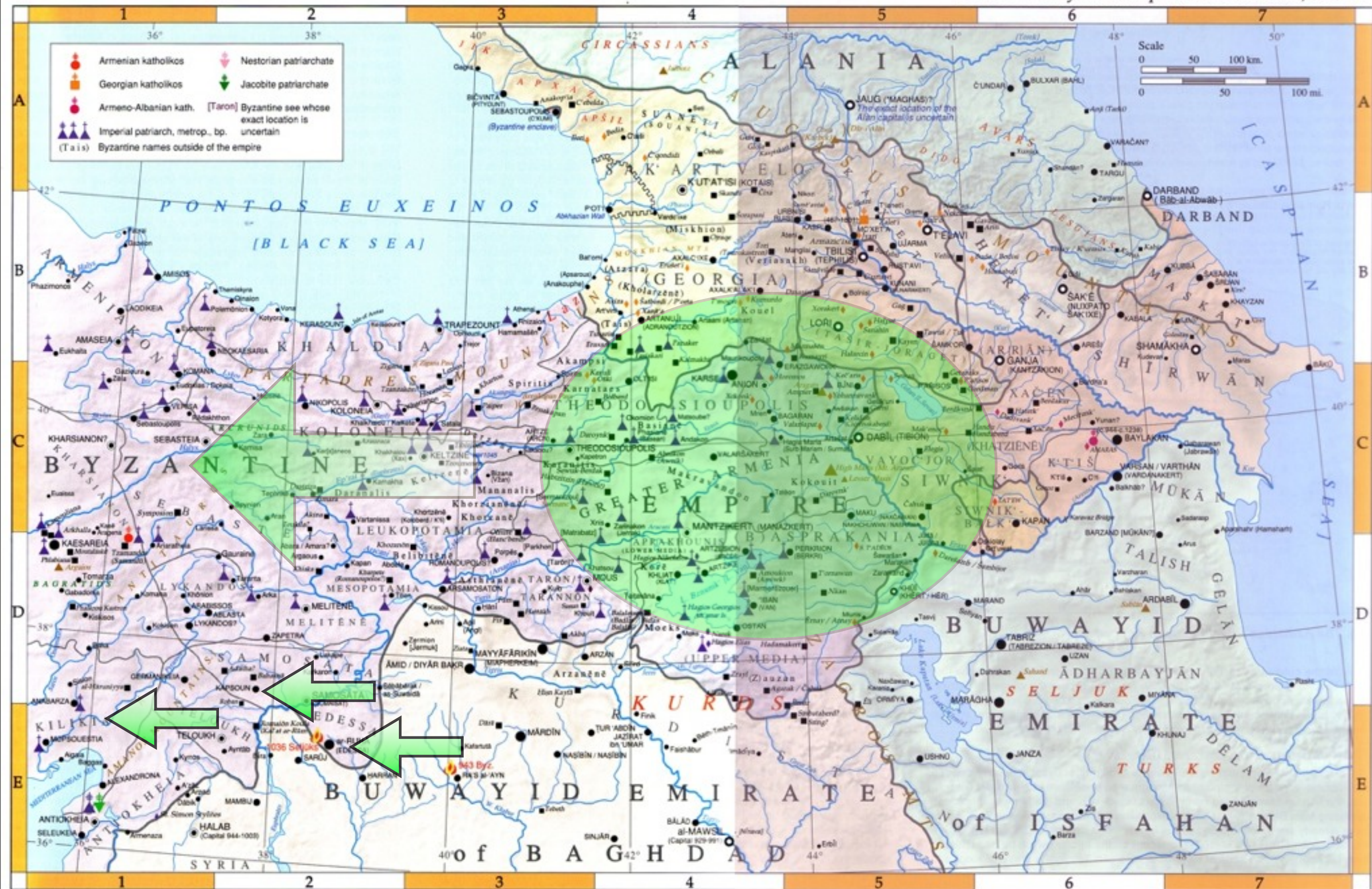


105. *The Third Byzantine Expansion into Armenia, 966–1064*



105. *The Third Byzantine Expansion into Armenia, 966–1064*



105. *The Third Byzantine Expansion into Armenia, 966–1064*



# Making an edition





SCRIPTORIUM MONK AT WORK. (From *Lacroix*.)

101010

000 | 0 |

1 0 0 1 0 1

# Surviving manuscripts



# Surviving manuscripts

- Oldest full manuscript is Venice 887

# Surviving manuscripts

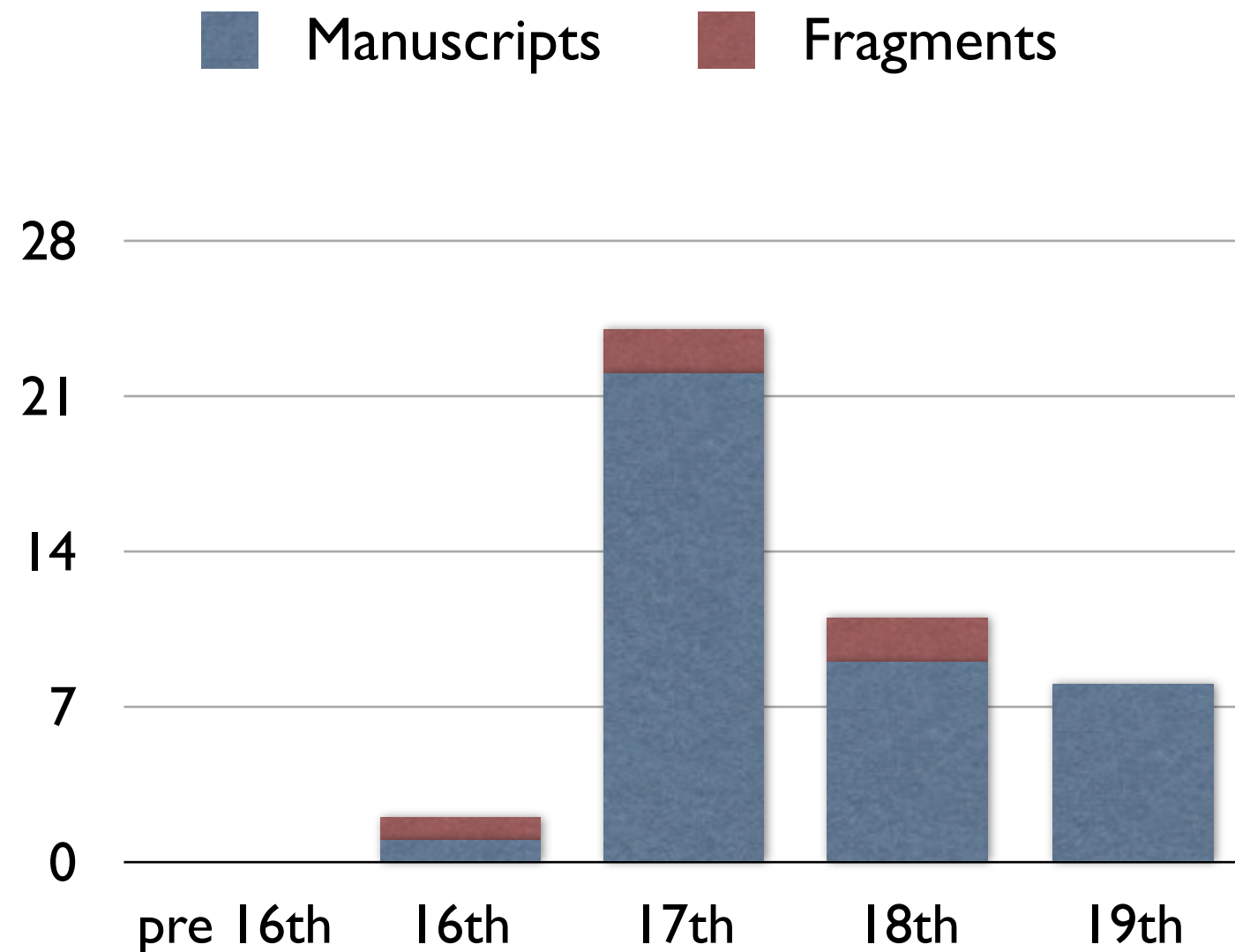
- Oldest full manuscript is Venice 887
- Next oldest is Wien 574



# Surviving manuscripts

- Oldest full manuscript is Venice 887
- Next oldest is Wien 574
- 24 of 43 (4 of 6 fragments) copied before 1700

# Extant manuscripts of the Chronicle



# Two manuscript groups

- Group 1: like V887
  - Text generally complete
  - Transmitted with the Life of St. Nerses (Mesrop the Priest)
- Group 2: like W574
  - Text truncated near the year 1096/7
  - Transmitted with specific long sequence of texts

# Matenadaran 1896

# Matenadaran 1896

- Copied in 1689 in Bitlis

# Matenadaran 1896

- Copied in 1689 in Bitlis
- Uniquely preserves two passages of text

# Matenadaran 1896

- Copied in 1689 in Bitlis
- Uniquely preserves two passages of text
- These lacunae known to other copyists

# Matenadaran 1896

- Copied in 1689 in Bitlis
- Uniquely preserves two passages of text
- These lacunae known to other copyists
- But lots of manuscripts are older. Hm.



# Making the edition

# Making the edition

- Transcription

# Making the edition

- Transcription
- Collation & text analysis

# Making the edition

- Transcription
- Collation & text analysis
- Editing the text

# Making the edition

- Transcription
- Collation & text analysis
- Editing the text
- Publication

# Collection



















(great conference)





(thanks domm!)



**(thanks vienna.pm!)**



(thanks again!)

# Transcription

# Transcription

- The most time-consuming part

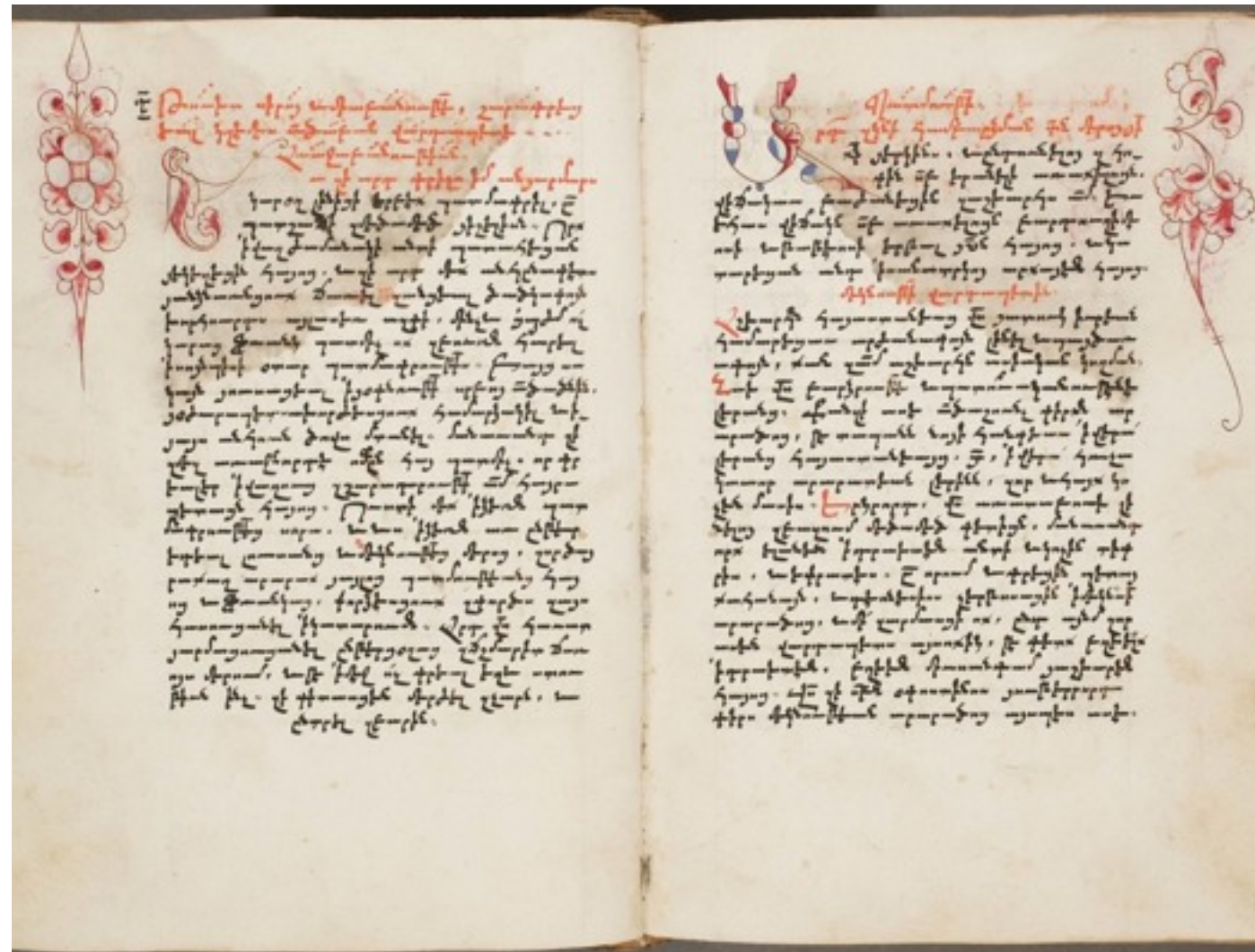


# Transcription

- The most time-consuming part
- No OCR for manuscripts, yet

# Transcription

- The most time-consuming part
- No OCR for manuscripts, yet
- (Any OCR experts? Talk to me.)



# Start with a manuscript

# Into plain text

դ. Թուխտ սիրոյ և միաբանութեան, շարագրեցեալ կղէմէս աստուածաբան վարդապետէ:

Առաջաբանութիւն:

Նա զի արդ գրիչս իմ անյարմարս կարօղ լինիցի երբէկ պատմագրիլ, ըստ պատշաճի զմեծամեծ յիշելիսն: Որք ի վաղ ժամանակի անտի պատահեցան յեկեղեցին հայոց. և զի արդ մեք անհրնազէտս յանձնառնցուք ճառել զանցեալ ծածկագոյն խորհուրդս ալլասեռ ազգի, մինչև ցայժմ ո՛չ կարաց Փրանկ պատմիչ ոք զբուռն հարել ի սոյնպիսի օտար պատմագրութիւնս: Բայց սակայն յուսացեալ ի յօգնութիւն սրբոյ աստուածածնին յօժարապէտս ախորժեսցուք համարձակիլ և ի յայս անհոռն ծովս մտանել...

# XML solution: TEI

```
<!-- ... -->
<div n="4">
  <head><hi rend="red">Թուխտ սիրոյ և միաբանութե<ex>ան</ex>, շարագրեց<lb/>
    եալ կղէմէս ա<ex>ստուա</ex>ծաբան վարդապետէ: <lb/>
    Առաջաբանութիւն: </hi>
  </head>
  <p><hi rend="ornament">Ն</hi><hi rend="red">ա զի արդ գրիչս իմ անյարմարս</hi> <lb/>
    կարօղ լինիցի երբէկ պատմագրիլ, <expan>ըստ</expan> <lb/>
    պատշաճի զմեծամեծ յիշելիսն: Որք <lb/>
    ի վաղ ժամանակի անտի պատահեցան <lb/>
    յեկեղեցին հայոց. և զի արդ մեք անհընագէտս <lb/>
    յանձնառնցուք ճառել զանցեալ ծածկագոյն <lb/>
    խորհուրդս ալլասեռ ազգի, մինչև ցայժմ ո՛չ <lb/>
    կարաց Ֆրանկ պատմիչ ոք զբուռն հարել <lb/>
    ի սոյնպիսի օտար պատմագրութի<ex>ւն</ex>ս: Բայց սա<lb/>
    կայն յուսացեալ ի յօգնութի<ex>ւն</ex> սրբոյ ա<ex>ստուա</ex>ծածնին <lb/>
    յօժարապէտս ախորժեսցուք համարձակիլ և ի <lb/>
    յայս անհոռն ծովս մտանել: ... <lb/>
  </p>
</div>
<!-- ... -->
```





Ի Թուիս սիրոյ ամուսնութի, շարժարեւոյ  
եւ ի զիս ամուսնութի զարդարեալս - -  
Համբարձնութիւն  
- ի - րդ - ԳԵԼԵ ԻՆ - անդամս  
Կ-րօղ լիկցի երեւոյ զարմարեւ. Ե  
զարշապ շեծ-ժծ յիշեալս. Որ  
ի վարժ-մանիւ անտ զարմարեւ  
յիշեալցի հոյոյ. Եւ ի րդ մա անհր-մար  
յանք-անցաւ ծառի զանցեալ ծածկ-պոյն  
խորհուրդս անստ-պիտ, մեւ ի յոյժ ու  
ի-րոյ Զաւան զարմարեւ ու զհոս անհր-մար  
իսոյնիս օրոյ զարմարեւ. Բայց  
կոյն յոս-յեւ ի յօգնութի արեւոյ ամ-ծախ.  
յօգն-արեւ-իսոյնիս հոս-իսիւ Եւ  
յոյ անհոս ծոպս մարեւ. Եւ ան-անտ ի  
զիւ ան-անտ ծոպս հոյ զարմարեւ. ու ի  
եւ ի վարժ-մանիւ զարմարեւ. Եւ հոյ-  
արեւ-անտ հոյ. Որ-անտ մա ի իսոյն զարմարեւ.

# Better TEI XML text

```
<!-- ... -->
<div n="4">
  <head><w>Թուխտ</w> <w>սիրոյ</w> <w>և</w> <w>միաբանու<ex>թեան</ex>,</w> <w>շարագրեց<lb/>եալ</w> <w>կղէմէս</w>
<w><ex>աստուած</ex>աբան</w> <w>վ
արդապետէ:</w><lb/>
  <w>Առաջաբանութիւն:
</head>

<p><hi><w>Նա</w> <w>զի</w> <w>արդ</w> <w>գրիչս</w> <w>իմ</w> <w>անյարմարս</w></hi> <lb/>
  <w>կարող</w> <w>լինիցի</w> <w>երբէկ</w> <w>պատմագրիլ,</w> <w><expan>ըստ</expan></w> <lb/>
  <w>պատշաճի</w> <w>զմեծամեծ</w> <w>յիշելիսն:</w> <w>Որք</w> <lb/>
  <w>ի</w> <w>վաղ</w> <w>ժամանակի</w> <w>անտի</w> <w>պատանեցան</w> <lb/>
  <w>յեկեղեցին</w> <w>հայոց.</w> <w>և</w> <w>զի</w> <w>արդ</w> <w>մեք</w> <w>անհրնագէտս</w> <lb/>
  <w>յանձնառնցութ</w> <w>հառել</w> <w>զանցեալ</w> <w>ծածկագոյն</w> <lb/>
  <w>խորհուրդս</w> <w>այլասեռ</w> <w>ազգի,</w> <w>մինչև</w> <w>ցայժմ ր՛չ</w> <lb/>
  <w>կարաց</w> <w>Ֆրանկ</w> <w>պատմիչ</w> <w>ոք</w> <w>գրուոն</w> <w>հարել</w> <lb/>
  <w>ի</w> <w>սոյնպիսի</w> <w>օտար</w> <w>պատմագրու<ex>թիւնս</ex>:</w> <w>Բայց</w> <w>սա<lb/>կայն</w> <w>յուսացեալ</
w> <w>ի յօգնու<ex>թիւն</e
x></w> <w>սրբոյ</w> <w><ex>աստուած</ex>ածնին</w> <lb/>
  <w>յօժարապէտս</w> <w>ախորժեսցութ</w> <w>համարձակիլ</w> <w>և</w> <w>ի</w> <lb/>
  <w>յայս</w> <w>անհոռն</w> <w>ծովս</w> <w>մտանել:</w> ... <lb/>
</p>
</div>
<!-- ... -->
```

# Perl to the rescue #1

- XML is a terrible thing to edit



# Perl to the rescue #1

- XML is a terrible thing to edit
- I want a transcription markup that I can convert to TEI XML later

# Perl to the rescue #1

- XML is a terrible thing to edit
- I want a transcription markup that I can convert to TEI XML later
- Obviously the world needs yet another markup syntax

<seg type="word"><subst><del>ໄພງຸ່ງ</del><add  
place="overwrite">ໄຫຫຫມຸ່ງ</add></subst>

$\pm$ - $\text{L}_{\text{WJNG}}^+$ (overwrite) $\text{L}_{\text{NNNG}}^+$

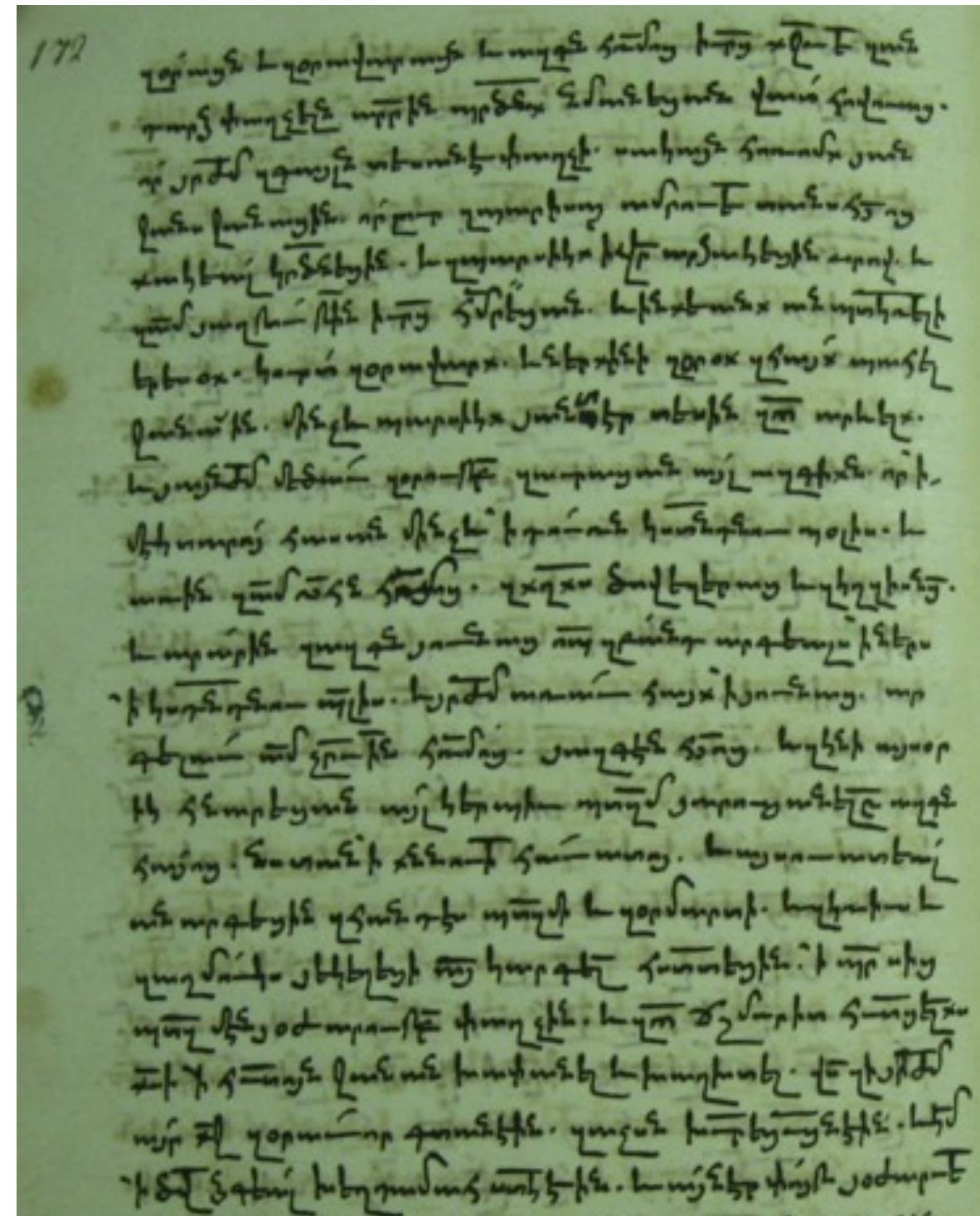
# TEI markup

[172]

զօ՛րացն և զօրավարացն և ազգն հ\ո\ո\ո\մոց իւր\ո\ց ք\ա\ջ\ո\ւ\թ\ե\ան\ զան  
դարձ փաղչելն ար\ա\րիւն պ\ա\րծ\ա\նք նմանեացն վատ՝ հովուաց,  
ո՛ր յ\ո\րժ\ա\մ զգալն տեսանէ փաղչի, սակայն հոռոմք յան  
ջանս ջանացին, ո՛ր լուր զպարիսպ ամրու\թ\ե\ան\ տանս հ\ա\յոց  
քակեա՛ւ ի կ\ո\րծ\ա\նեցին, և զպարսիկք ի վ\ե\ր\ա\յ\ արձակեցին սրով, և  
զամենայն\ յաղթո՛ւ թիւնն իւր\ո\ց հ\ա\մ\ա\րեցան, և ինքեանք անպ\ա\տկ\ա\ռելի  
երեսօք, կուրտ՝ զօրավարք, և ներքինի զօրօք զհայ՝ ք պահել  
ջանա՛յսին, մինչև պարսիկք յանտէր տեսին զամենայն արևելք.  
և յայնժ\ա\մ մեծա՛ւ զօրութեամբ զօրացնն այ՛ լազգիքն, որ ի  
մէկ տարո՛յս հասան մինչև ի դո՛ւռն կ\ո\ստ\ա\նդ\նուպօլիս, և  
առին զամենայն\ լաշխարհն և իւրեքս (overwrite)\ո\ո\ո\մ\տոց, զք\ա\ղ\ա\քս  
ծովեզերաց և զկղզիս նոց\ա\,

և արա՛րին զազգն յունաց ո՛ր\ա\յ\էս\ զբա՛նդարգեալս ի ներս  
ի կ\ո\ստ\ա\նդ\նուպօլիս. և յ\ո\րժ\ա\մ առա՛ւ հայք ի յունաց, ար  
գելա՛ւ ամենայն չ\ա\ր\ո\ւ\թ\իւնն հոռոմո՛ց, յազգէն հ\ա\յոց, և զկնի այսօր  
իկ հնարեցան այ՛ ի կերպիւ պ\ա\տ\երա\զմ յարուցանեալ ^ընդ^ ազգն  
հա՛յոց, նստան ի քննութ\իւնն հա՛ւատոյ, և այսու ատեա՛ւ ի  
անարգեցին զհանդէս պ\ա\տ\երա\զմի և զօրմարտի, և զկռիւս և  
զաղմո՛ւ կս յեկեղեցի անստուծո՛յ կարգեալ հ\ա\ստ\ա\տեցին. ի պ\ա\րսից  
պ\ա\տ\երա\զմէն յօժարութեամբ փախչին, և զամենայն\ ճշմարիտ

հ\ա\ւ\ա\տ\ա\ցեալքս  
ք\րիստոս\ի ի հ\ա\ւ\ա\տոյն ջանան խափանել և խաղիտել, վ\ա\ս\ն\ զի յ\ո\րժ\ա\մ  
այր ք\ա\ջ զօրա՛ւ որ գտանէին, զաչսն խ\ա\ւ\ա\րեց\ո\ւց\ա\նէին, և կ\ա\մ  
ի ծ\ո\վ ձգեա՛ւ ի բեղդամահ ս\ա\տ\ա\կէին. և այ՛ն էր փո՛յս յօժար\ո\ւ\թ\ե\ան\



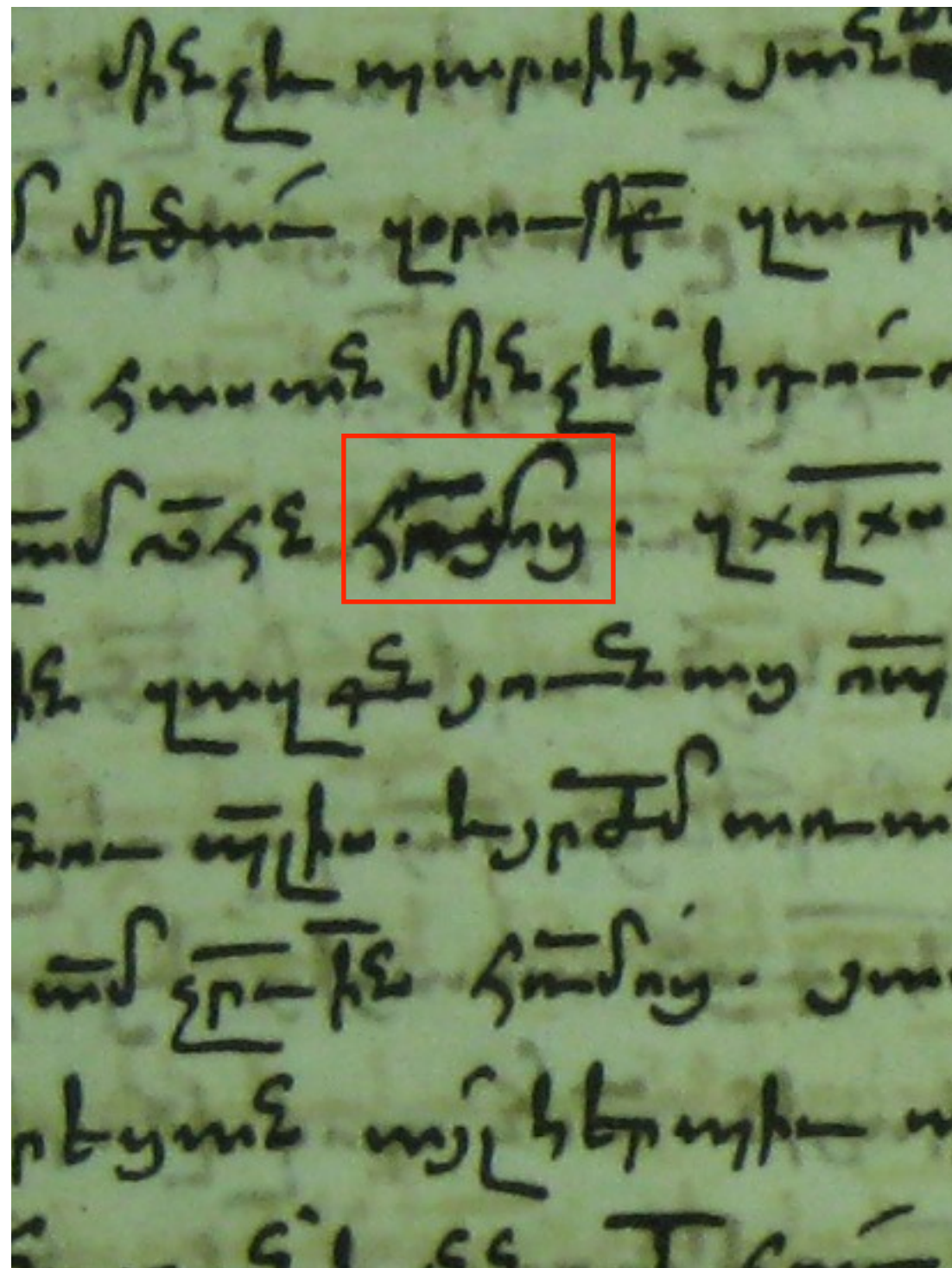


# TEI markup

ⲕⲓⲧⲉ-ⲙⲓⲛⲓⲧⲉ(overwrite)\ⲛ\ⲛ\ⲛ\ⲙⲓⲧⲉⲛⲓⲛⲓⲛⲓ,

---

```
<seg type="word">ⲕⲓⲧⲉ  
<subst><del>ⲙⲓⲛⲓⲧⲉ</del>  
  <add place="overwrite">  
    <ex resp="#tla">ⲛ</ex>ⲛ  
    <ex resp="#tla">ⲛ</ex>ⲙⲓⲧⲉ</add>  
</subst>ⲛⲓⲛⲓⲛⲓ,</seg>
```



# TEI markup

Հ\nռ\nռ\nռ\nմոց իւր\nռ\nց ք\nւ\nա\nչ\nռ\nւ\nթ\nե\nւ\nա\nն\nւ

<seg type="word">Հ

<ex>ռ</ex>ռ

<ex>ռ</ex>մոց</seg>

<seg type="word">իւր

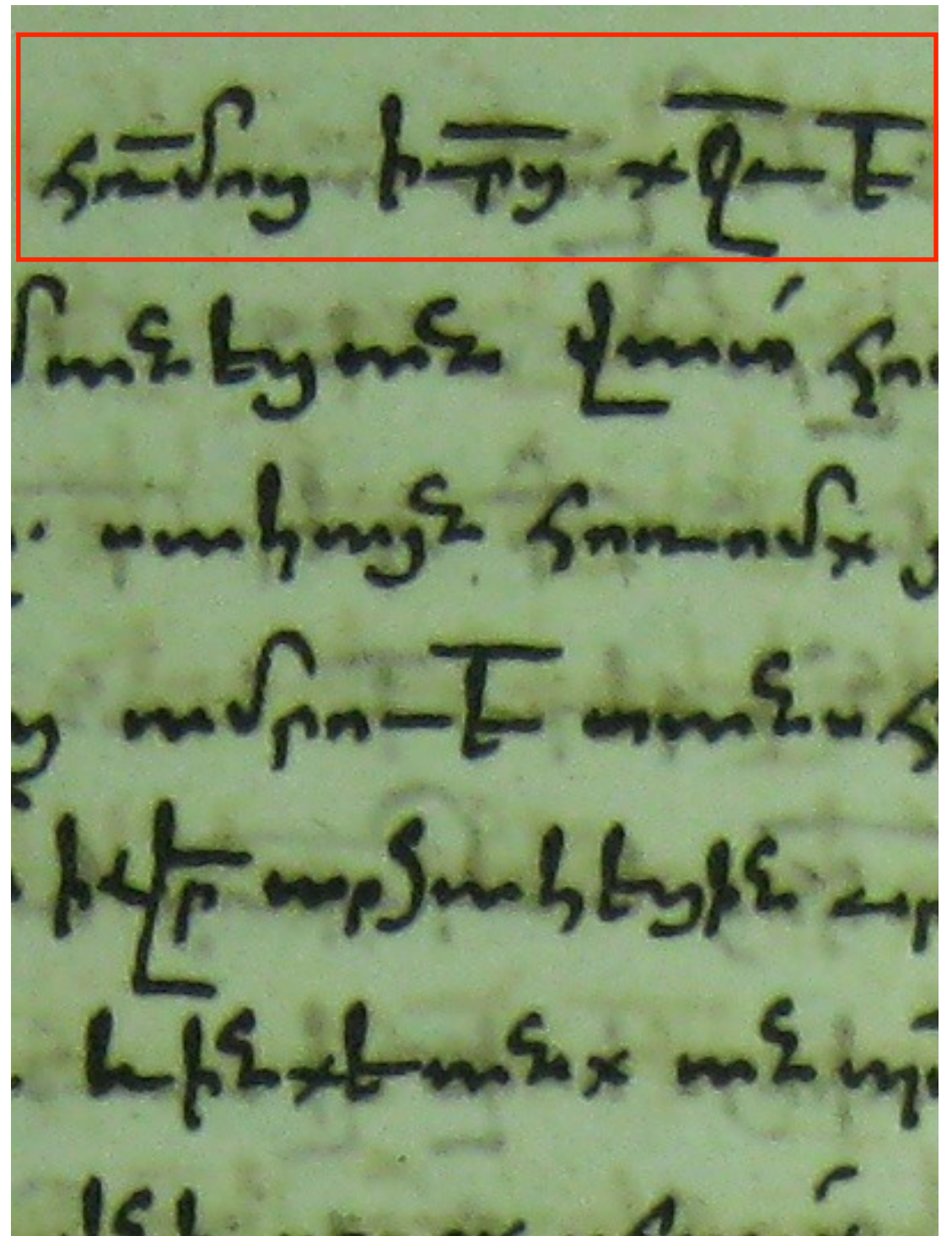
<ex>ռ</ex>ց</seg>

<seg type="word">ք

<ex>ւ</ex>չ<ex>ռ</ex>ւ<ex>թ</

ex>ե<ex>ւ\nւ</ex>

</seg>



# TEI XML text

```
<!-- ... -->
<div n="4">
  <head><w>Թուխտ</w> <w>սիրոյ</w> <w>և</w> <w>միաբանու<ex>թեան</ex>,</w> <w>շարագրեց</w>եալ</w>
<w>կղէմէս</w> <w><ex>աստուած</ex>աբան</w> <w>վ
արդապետէ:</w></b>
  <w>Առաջաբանութիւն:
</head>

<p><hi><w>Նա</w> <w>զի</w> <w>արդ</w> <w>գրիչս</w> <w>իմ</w> <w>անյարմարս</w></hi> </b>
<w>կարող</w> <w>լինիցի</w> <w>երբէկ</w> <w>պատմագրիլ,</w> <w><expan>ըստ</expan></w> </b>
<w>պատշաճի</w> <w>զմեծամեծ</w> <w>յիշելիսն:</w> <w>Որք</w> </b>
<w>ի</w> <w>վաղ</w> <w>ժամանակի</w> <w>անտի</w> <w>պատահեցան</w> </b>
<w>յեկեղեցին</w> <w>հայոց.</w> <w>և</w> <w>զի</w> <w>արդ</w> <w>մեք</w> <w>անհրնագէտս</w> </b>
<w>յանձնառնցութ</w> <w>հառել</w> <w>զանցեալ</w> <w>ծածկագոյն</w> </b>
<w>խորհուրդս</w> <w>ալլասեռ</w> <w>ազգի,</w> <w>մինչև</w> <w>ցայժմ ո՛չ</w> </b>
<w>կարաց</w> <w>Ֆրանկ</w> <w>պատմիչ</w> <w>ոք</w> <w>զբուռն</w> <w>հարել</w> </b>
<w>ի</w> <w>սոյնպիսի</w> <w>օտար</w> <w>պատմագրու<ex>թիւնս</ex>:</w> <w>Բայց</w> <w>սա</b>
<w>կայն</w> <w>յուսացեալ</w> <w>ի յօգնու<ex>թիւն</ex>
<w>սրբոյ</w> <w><ex>աստուած</ex>ածնին</w> </b>
  <w>յօժարապէտս</w> <w>ախորժեսցութ</w> <w>համարձակիլ</w> <w>և</w> <w>ի</w> </b>
  <w>յայս</w> <w>անհոռն</w> <w>ծովս</w> <w>մտանել:</w> ... </b>
</p>
</div>
<!-- ... -->
```



# Collation

"The collation of manuscripts requires the infuriating accuracy of a pedant and the obsessive stamina of an idiot. It is therefore an ideal task for a computer."

—Peter Robinson, "Collation and Textual Criticism", *LLC* vol. 4 no. 2, 1989

“If nothing else, the choice of programming language was easy. [redacted] was developed especially for text processing, and so is very rich in text-handling functions.

“It is, for instance, quite easy to use [redacted] to ask the computer to 'read the first word in line x; compare this word with the first word in line y; if a match is found do a; if not, do b'. As that—greatly elaborated—is essentially what I wanted to do, [redacted] was the obvious language of choice.”

—Peter Robinson, “Collation and Textual Criticism”, *LLC* vol. 4 no. 2, 1989

“If nothing else, the choice of programming language was easy. SNOBOL was developed especially for text processing, and so is very rich in text-handling functions.

“It is, for instance, quite easy to use SNOBOL to ask the computer to 'read the first word in line x; compare this word with the first word in line y; if a match is found do a; if not, do b'. As that—greatly elaborated—is essentially what I wanted to do, SNOBOL was the obvious language of choice.”

—Peter Robinson, “Collation and Textual Criticism”, *LLC* vol. 4 no. 2, 1989



# Collation

- My collation tool is in Perl

# Collation

- My collation tool is in Perl
- Uses `Algorithm::Diff`

# Collation

- My collation tool is in Perl
- Uses `Algorithm::Diff`
- Uses `Text::WagnerFischer`



# Collation

- My collation tool is in Perl
- Uses `Algorithm::Diff`
- Uses `Text::WagnerFischer`
  - ... `::Armenian`

# Collation

# Collation

- need to align words with each other



# Collation

- need to align words with each other
- ...across many manuscripts

# Collation

- need to align words with each other
- ...across many manuscripts
- ...even when the words aren't exactly the same  
(e.g. “յաշխարհի” vs. “աշխարհն”)



```

<!-- ... -->
<p>
  <w><app>
    <rdg wit="#A #C">յալսմ</rdg>
    <rdg wit="#B #D #E">ալս</rdg>
  </app></w>
  <w><app>
    <rdg wit="#A #B #D #E">ամենալն</rdg>
    <rdg wit="#C">ամի</rdg>
  </app></w>
  <w><app>
    <rdg wit="#A">եղեկոյն</rdg>
    <rdg wit="#B #C #D">եղեկոյ</rdg>
    <rdg wit="#E">եղեկոյս</rdg>
  </app></w>
  <w>նստուցանեն</w>
  <w type="prefix"><app>
    <rdg wit="#A #B #D #E">դ</rdg>
    <rdg wit="#C">յ</rdg>
  </app></w>
  <w>աթոռ</w>
  <w>հայրապետութեան</w>
  <w><app>
    <rdg wit="#A #C #D">ի</rdg>
  </app></w>
  <!-- ... -->
</p>
<!-- ... -- >

```



```

<!-- ... -->
<p>
  <w><app>
    <rdg wit="#A #C">յալսմ</rdg>
    <lem wit="#B #D #E">ալս</rdg>
  </app></w>
  <w><app>
    <lem wit="#A #B #D #E">ամենալն</rdg>
    <rdg wit="#C">ամի</rdg>
  </app></w>
  <w><app>
    <lem wit="#A">եղեկոյն</rdg>
    <rdg wit="#B #C #D">եղեկոյ</rdg>
    <rdg wit="#E">եղեկոյս</rdg>
  </app></w>
  <w>նստուցանեն</w>
  <w type="prefix"><app>
    <rdg wit="#A #B #D #E">դ</rdg>
    <lem wit="#C">յ</rdg>
  </app></w>
  <w>աթոռ</w>
  <w>հայրապետութեան</w>
  <w><app>
    <lem wit="#A #C #D">ի</rdg>
  </app></w>
  <!-- ... -->
</p>
<!-- ... -->

```

# Our text & apparatus

1 ասս ամենայն եղեւոցն նստուցանեն զաթոռ  
ի թաւթուր,

1 ասս] յասսմ AC 1 ամենայն] ամի C 1 եղեւոցն] եղեւոց BDE եղեւոցս C  
1 զաթոռ] յաթոռ C 2 ի թաւթուր] om. BE

...

```

<!-- ... -->
<p>
  <w><app>
    <rdg wit="#A #C">յալսմ</rdg>
    <lem wit="#B #D #E">ալս</rdg>
  </app></w>
  <w><app>
    <lem wit="#A #B #D #E">ամենալն</rdg>
    <rdg wit="#C">ամի</rdg>
  </app></w>
  <w><app>
    <lem wit="#A">եղեկոյն</rdg>
    <rdg wit="#B #C #D">եղեկոյ</rdg>
    <rdg wit="#E">եղեկոյս</rdg>
  </app></w>
  <w>նստուցանեն</w>
  <w type="prefix"><app>
    <rdg wit="#A #B #D #E">դ</rdg>
    <lem wit="#C">յ</rdg>
  </app></w>
  <w>աթոռ</w>
  <w>հայրապետութեան</w>
  <w><app>
    <lem wit="#A #C #D">ի</rdg>
  </app></w>
  <!-- ... -->
</p>
<!-- ... -->

```

```

<!-- ... -->
<p>
  <w><app>
    <lem wit="#A #C">յայտմ</rdg>
    <rdg wit="#B #D #E">այս</rdg>
  </app></w>
  <w><app>
    <rdg wit="#A #B #D #E">ամենայն</rdg>
    <lem wit="#C">սփի</rdg>
  </app></w>
  <w><app>
    <lem wit="#A">եղեկոյն</rdg>
    <rdg wit="#B #C #D">եղեկոյ</rdg>
    <rdg wit="#E">եղեկոյս</rdg>
  </app></w>
  <w>նստուցանեն</w>
  <w type="prefix"><app>
    <rdg wit="#A #B #D #E">դ</rdg>
    <lem wit="#C">յ</rdg>
  </app></w>
  <w>սթոռ</w>
  <w>հայրապետութեան</w>
  <w><app>
    <lem wit="#A #C #D">ի</rdg>
  </app></w>
  <!-- ... -->
</p>
<!-- ... -- >

```



# New text & apparatus

1 յայսմ ամի եղեւոցն նստուցանեն զաթոռ  
ի թաւթուր,

1 յայսմ] այս BDE 1 ամի] ամենայն ABDE 1 եղեւոցն] եղեւոց BDE եղեւոցս C  
1 զաթոռ] յաթոռ C 2 ի թաւթուր] om. BE  
...

Word variation! Context:

վայրս բազմաչան և աշխատաւոր քննութեամբ գտեալ գրեցաք ի  
զշարագրական գրեալս >> զհարիւրից ամաց, զորս ի << բազում ժամանակաց  
հետաքննեալ հասու եղաք: Ընդ այնքանեաց տեսողացն և լսողացն,

Base զհարիւրից ամաց, զորս ի

----

E: հարիւրից ամացն զորս

B: զհարիւրից ամաց զորս ի

AD: զհարիւրից ամաց զորս ի

C: զհարիւրից ամաց զոր ի

=>

# Manuscript stemmas: the family tree

# Stemma construction

- Better stemma through analysis of collation results



# Stemma construction

- Better stemma through analysis of collation results
- Borrows statistical models from evolutionary biology

# Stemma construction

- Better stemma through analysis of collation results
- Borrows statistical models from evolutionary biology
- “Maximum parsimony” based upon DNA of specimens

# Stemma construction

- Better stemma through analysis of collation results
- Borrows statistical models from evolutionary biology
- “Maximum parsimony” based upon DNA of specimens
- Manuscripts are specimens

# Stemma construction

- Better stemma through analysis of collation results
- Borrows statistical models from evolutionary biology
- “Maximum parsimony” based upon DNA of specimens
- Manuscripts are specimens
- Biologists have DNA sequences; we have words.



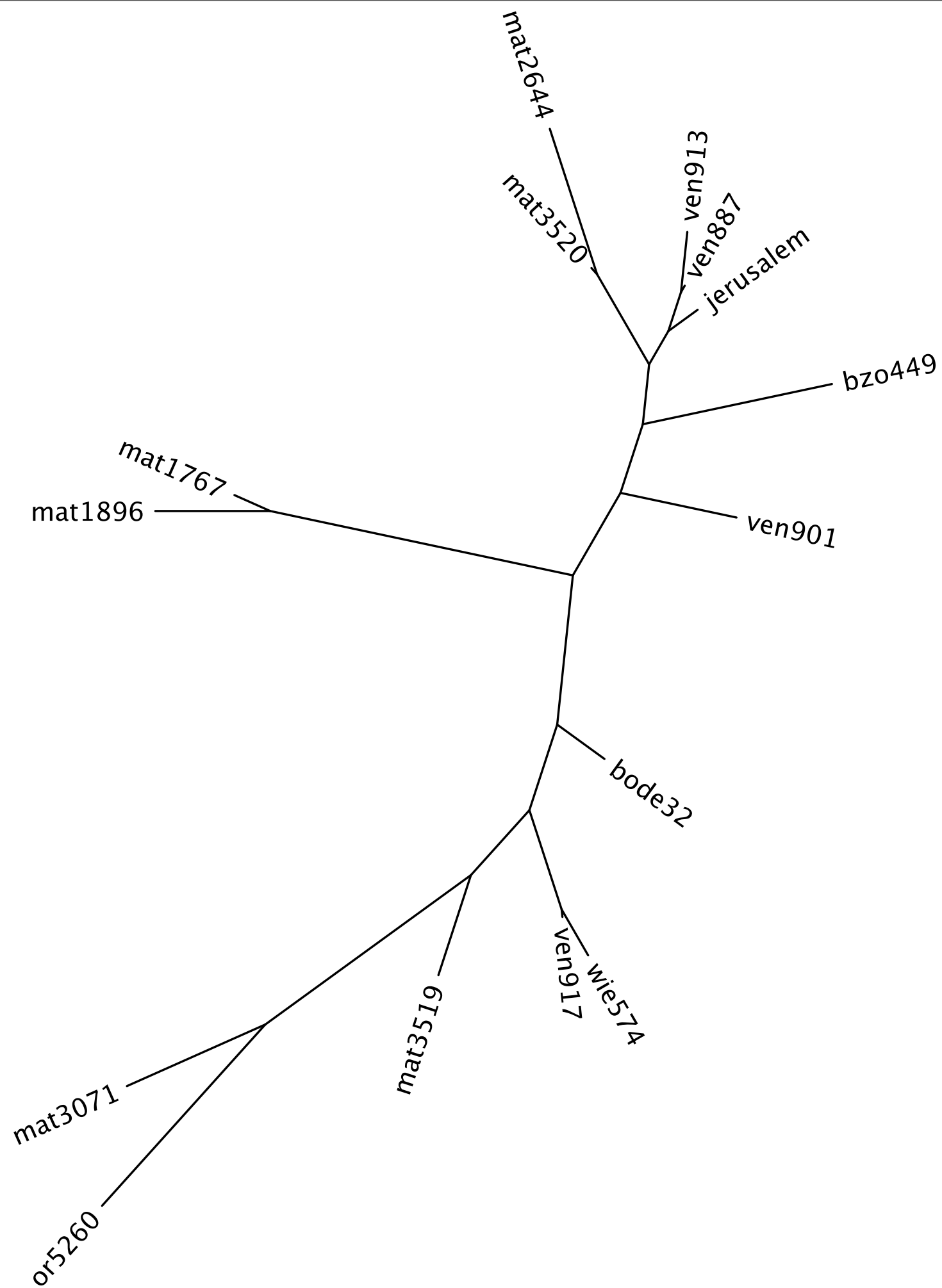


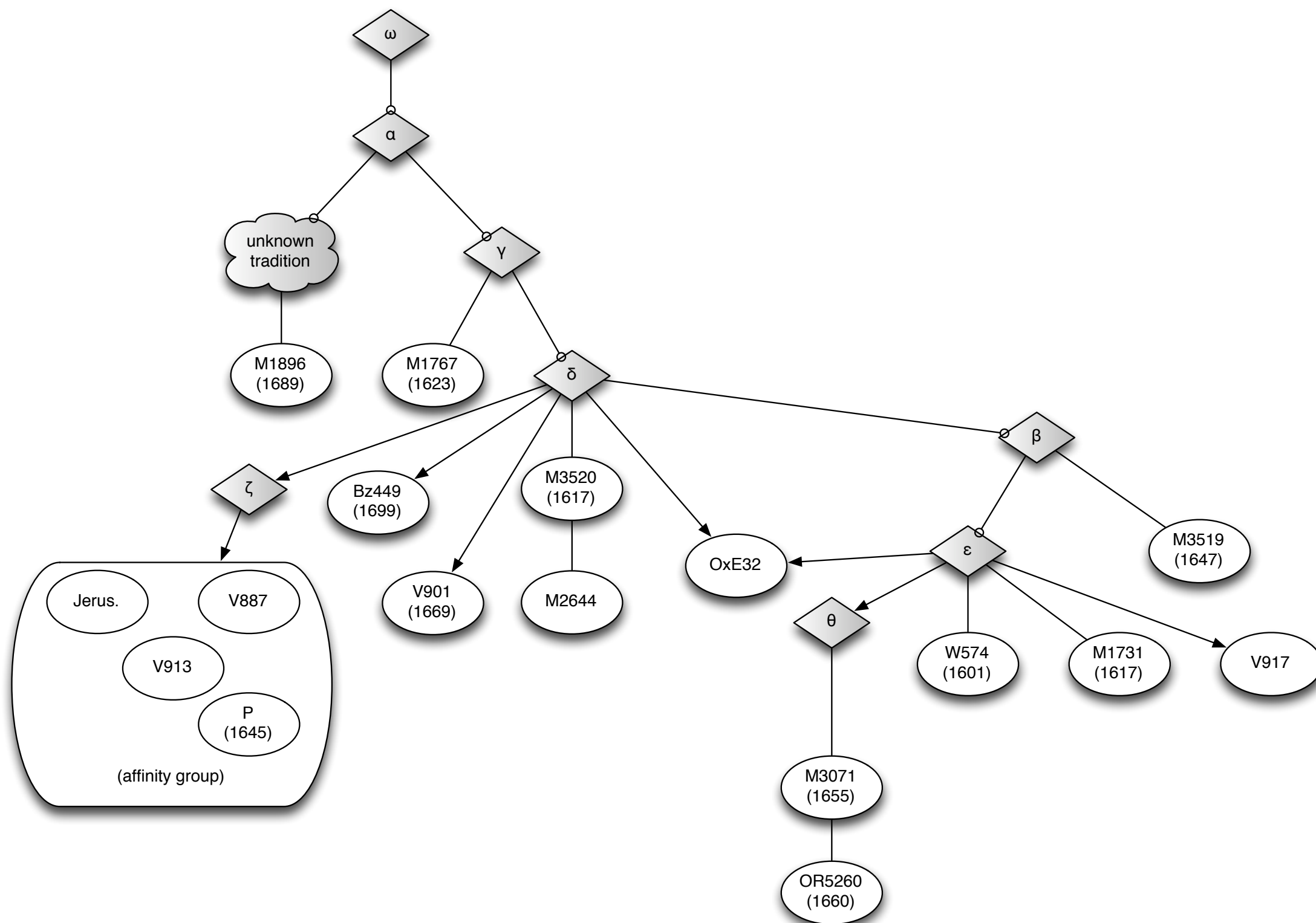
[illegible][illegible]

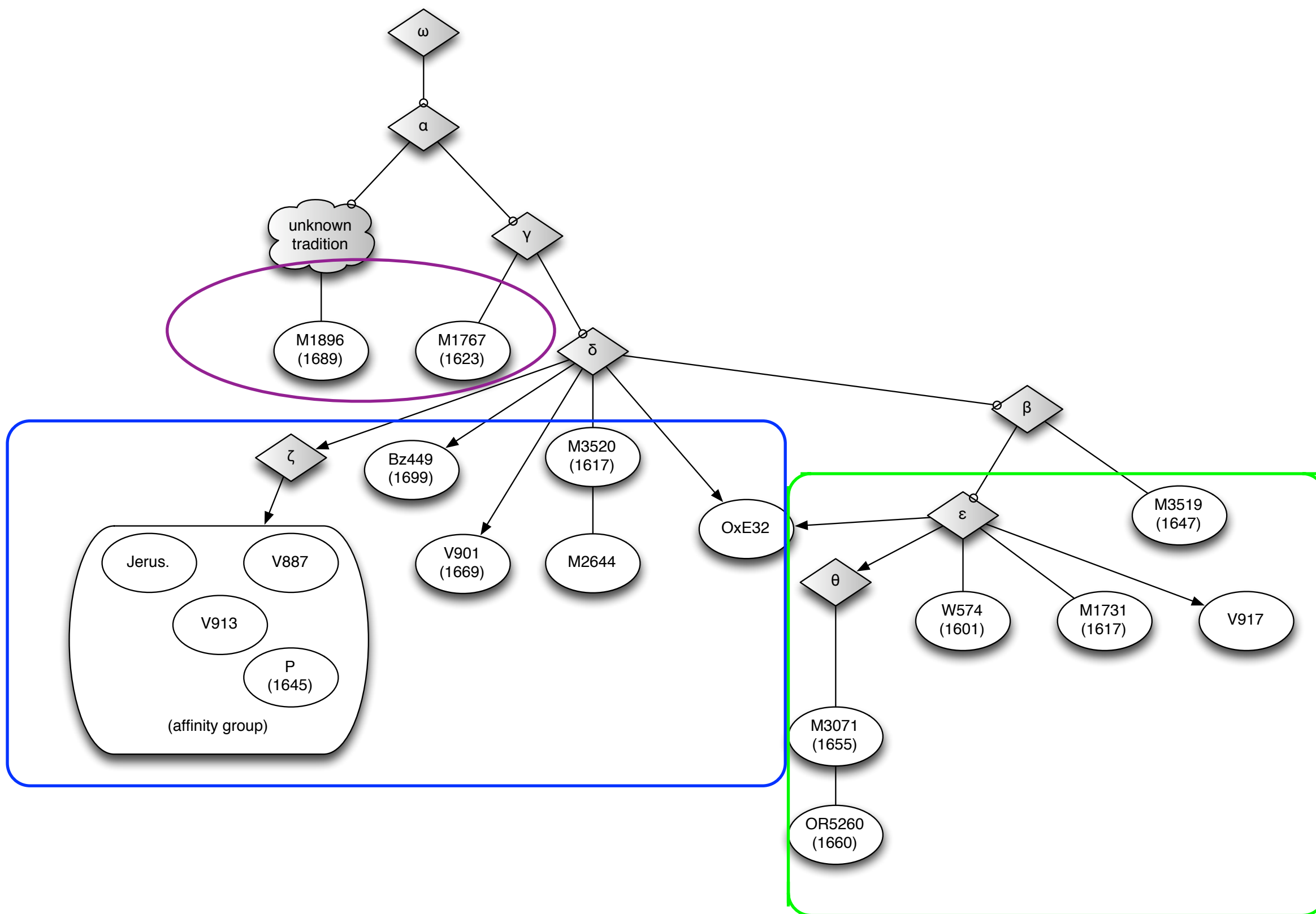
A  
B  
B  
A  
B  
A  
A  
B  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
B  
A  
B

[illegible]

B  
A  
C  
A  
A  
A  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
B  
A









# Vaporware

# Edition creation

- You've seen my poor excuse for a UI

# Edition creation

- You've seen my poor excuse for a UI
- Basic functionality is there

# Edition creation

- You've seen my poor excuse for a UI
- Basic functionality is there
- Need ability to add manuscripts over time

# Edition creation

- You've seen my poor excuse for a UI
- Basic functionality is there
- Need ability to add manuscripts over time
- Want to add hook for morphological analysis



# Publication

# Publication

- Final edition is in TEI XML format

# Publication

- Final edition is in TEI XML format
- Transforms exist for TEI => LaTeX

# Publication

- Final edition is in TEI XML format
- Transforms exist for TEI => LaTeX
- Can bypass word processor and text editor for print edition

# Other publication



# Other publication

- XML transforms also exist for online publication

# Other publication

- XML transforms also exist for online publication
- This gives:
  - searchable text

# Other publication

- XML transforms also exist for online publication
- This gives:
  - searchable text
  - easy updates

# Other publication

- XML transforms also exist for online publication
- This gives:
  - searchable text
  - easy updates
  - configurable set variants

# Other publication

- XML transforms also exist for online publication
- This gives:
  - searchable text
  - easy updates
  - configurable set variants
  - links to manuscript images where available



Thank you  
Danke schön  
Dakujem  
Հնարակալիկութիւն